Metabolic Engineering 47 (2018) 134–142

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

Glyco-Mapper: A Chinese hamster ovary (CHO) genome-specific glycosylation prediction tool

Benjamin G. Kremkow[a,b], Kelvin H. Lee[a,b,*]

[a] Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE 19716, USA
[b] Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711, USA

ARTICLE INFO

Keywords:
Chinese hamster ovary (CHO) cells
Genomic modeling
Glycosylation reaction network
Glycoform
Biopharmaceutical process development

ABSTRACT

Glyco-Mapper is a novel systems biology product quality prediction tool created using a new framework termed: Discretized Reaction Network Modeling using Fuzzy Parameters (DReaM-zyP). Within Glyco-Mapper, users fix the nutrient feed composition and the glycosylation reaction fluxes to fit the model glycoform to the reference experimental glycoform, enabling cell-line specific glycoform predictions as a result of cell engineering strategies. Glyco-Mapper accurately predicts glycoforms associated with genetic alterations that result in the appearance or disappearance of one or more glycans with an accuracy, sensitivity, and specificity of 96%, 85%, and 97%, respectively, for publications between 1999 and 2014. The modeled glycoforms span a large range of glycoform engineering strategies, including the altered expression of glycosylation, nucleotide sugar transport, and metabolism genes, as well as an altered nutrient feeding strategy. A glycoprotein-producing CHO cell line reference glycoform was modeled and a novel Glyco-Mapper prediction was experimentally confirmed with an accuracy and specificity of 95% and 98%, respectively. Glyco-Mapper is a product quality prediction tool that provides a streamlined way to design host cell line genomes to achieve specific product quality attributes.

1. Introduction

Chinese hamster ovary (CHO) cells accounted for the production of $85 billion of the $154 billion in global biotherapeutic sales in 2015 (LaMerie, 2016) because CHO cells are capable of producing large amounts of protein, sustaining high viability, resisting viral infection, and mimicking human product quality attributes (Jayapal et al., 2007; Xu et al., 2011). The protein product quality largely influences the efficacy, half-life, and immunogenicity of the therapeutic protein (Walsh and Jefferis, 2006), all of which greatly affect the patient's clinical response. The 2013 biopharmaceutical pipeline contained 431 recombinant proteins and monoclonal antibodies (mAbs) in various phases of clinical development (ABRC, 2013). Understanding and predicting the therapeutic product quality in accordance with targeted modifications will be greatly beneficial to biopharmaceutical production and is critical to meet quality guidelines in accordance with FDA guidance (US DHHS FDA, 2015).

Several kinetic and data-dependent models that establish frameworks to quantify and model glycosylation have previously been published. The Umaña model (Umaña and Bailey, 1997) mathematically depicted the glycosyltransferase activity of 8 enzymes, 33 species, and 33 reactions. Krambeck and Betenbaugh (2005) expanded upon the Umaña model by incorporating more variables and models 11 enzymes, 7565 species, and 22,871 reactions, increasing the model's complexity. Krambeck et al. (2009) further broadened the model to incorporate 19 enzymes and more than 10,000 species to determine the likely enzyme concentrations by generating an optimized synthetic mass spectrum. Liu and Neelamegham (2014) further developed glycan mass spectra analysis to construct biochemical reaction networks and calculate the associated fluxes for both N- and O-glycosylation associated pathways, in addition to determining enzyme activities. Spahn et al. (2016) employed Markov chain modeling to mathematically calculate parameters to reproduce various glycoform distributions and did not require user-provided kinetic information. Despite the power of current analytical methods, kinetic model parameter estimation and validation is difficult to achieve for all relevant enzymes using current experimental techniques. Moreover, measurement of glycan stereoisomers and confirmation of a glycan's specific production pathway is not yet possible on a routine basis.

In contrast to detailed kinetic and data-driven models, genome-scale reconstructions can model biological processes using algebraic mass balance equations and have been successfully applied towards

Abbreviations: CCM, central carbon metabolism; CHO, Chinese hamster ovary; DReaM-zyP, Discretized Reaction Network Modeling using Fuzzy Parameters; EPO, erythropoietin; mAb, monoclonal antibody; SEAP, secreted alkaline phosphatase; siRNA, short interfering RNA

* Correspondence to: 15 Innovation Way, Newark, DE 19711, USA.
E-mail address: KHL@udel.edu (K.H. Lee).

https://doi.org/10.1016/j.ymben.2018.03.002
Received 17 May 2017; Received in revised form 7 December 2017; Accepted 1 March 2018
Available online 06 March 2018
1096-7176/

mammalian systems in many contexts (Duarte et al., 2007; Selvarasu et al., 2010; Shlomi et al., 2008) and are now being used to study CHO (Chen et al., 2012; Chowdhury et al., 2015; Selvarasu et al., 2012). The current availability of the CHO genome (Xu et al., 2011) offers an opportunity to investigate CHO-specific product quality using a genome scale reconstruction-based model. This type of model has reduced computational requirements compared to detailed kinetic models, but does not fully capture the variety of products generated by a non-template driven process. Here we report for the first time a modeling framework rooted in genome reconstruction but using reaction flux flow stoichiometry, discretized variable state parameters, and mass balances termed Discretized Reaction Network Modeling using Fuzzy Parameters (DReaM-zyP) to predict the likely therapeutic protein glycan patterns of the non-template driven glycosylation process. The specific CHO glycosylation DReaM-zyP-based tool (Glyco-Mapper) includes all CHO N-glycosylation genes, as well as nucleotide sugar synthesis, transporter, and glycosylation-relevant metabolism genes. Glyco-Mapper models and predicts the published cell-engineered glycoforms from 1999 to 2014 (Goh et al., 2014; Imai-Nishiya et al., 2007; Kanda et al., 2007; Malphettes et al., 2010; Maszczak-Seneczko et al., 2013; Naso et al., 2010; Onitsuka et al., 2012; Sealover et al., 2013; Tsukahara et al., 2006; Weikert et al., 1999) with an accuracy of 96% and is currently implemented in Microsoft Excel, which does not require the user to have extensive knowledge of modeling software or programming. The goal of Glyco-Mapper version 1.0 is to establish a simplified framework to provide useful glycoform pattern information, not to quantitatively predict various glycan concentrations, or levels, apart from their presence or absence.

2. Materials and methods

2.1. Gene and glycan composition

Glyco-Mapper contains 59 N-linked glycosylation genes (Supp. Table 1) and 92 metabolism-related genes (Supp. Table 2), describing the central carbon metabolism (CCM), nucleotide sugar synthesis, and nucleotide sugar transporter pathways (Supp. Fig. 1). The genes were manually verified to be present within the CHO and Chinese hamster genomes (Brinkrolf et al., 2013; Lewis et al., 2013; Xu et al., 2011) and the gene sequences were obtained from the 2014 RefSeq CHO-K1 and CH genome annotations (Hammond et al., 2012; Kremkow et al., 2015). Most of the N-linked glycosylation genes were obtained from Xu et al. (2011) and supplemental genes were added to the model from literature (Bosques et al., 2010). The N-linked glycosylation enzyme functions span the entire N-glycosylation reaction network and range from the production of the glycan intermediate in the endoplasmic reticulum to the degradation of the glycan outside of the Golgi. The metabolism-related genes were identified from a published CHO CCM model (Ahn and Antoniewicz, 2012), as were the genes involved with the nucleotide sugar production and nucleotide sugar transport pathways (Hills et al., 2001).

Both the KEGG database and literature were used to define the reactants, products, and enzymatic reaction conditions (Taniguchi et al., 2002; Kanehisa and Goto, 2000) for each enzyme coded within the CHO genome. However, the gene network was modeled using simplified reaction equations to generate glycans rather than comprehensive Michaelis-Menten kinetics. The Glyco-Mapper flux flow stoichiometry is based upon the definition of glycans by their monosaccharide components (Supp. Table 5), which glycosylation genes are required to produce each glycan (Supp. Table 6), the components each glycan requires of the CCM reaction network (Supp. Table 7), and glycosylation gene activities affecting the kinetic concentration of each specific glycan (Supp. Table 8). Glyco-Mapper models 448 non-stereospecific glycans (Supp. Table 5) representing more than 2600 distinct, stereospecific glycans. Stereoisomers are considered identical for this work because most current analytical glycan methods do not distinguish stereospecific glycans and there is no reported association between stereoisomers and biotherapeutic characteristics.

2.2. Experimental

A CHO-DUKX cell line expressing secreted alkaline phosphatase (SEAP) (Hayduk and Lee, 2005) was adapted to serum-free, suspension culture in 125 mL shake flasks (Corning, Oneonta, NY) containing 28 mL SFM4CHO medium (Hyclone Laboratories Inc., Logan, UT). The cells were cultured by routine passaging at 4 day intervals. Cultures were then seeded at $3{\cdot}10^5$ cells/mL and incubated with orbital agitation at 120 rpm in a 37 °C cell culture incubator with 5% CO_2 and 80% relative humidity. Cells were counted using a Countess II FL hemocytometer (ThermoFisher, Rockford, IL) with viability determined by the Trypan blue (Sigma-Aldrich, St. Louis, MO) exclusion method. The cells were harvested on day 3 and the supernatant was separated from the residual cells by centrifugation (180*g*, 6 min) and stored at − 20 °C until further use.

Supernatant samples were thawed simultaneously and filtered through a 0.22 μm filter (Millipore, Cork, Ireland). A SEAP-activity assay (ThermoFisher, Rockford, IL) was performed on all samples to quantify the SEAP protein concentration. SEAP was purified using a Reactive Green 19 pseudo-affinity chromatography column, generated according to the optimized protocol described by Ouyang et al. (2007). Briefly, Sepharose™ 6B (GE Healthcare, Uppsala, Sweden) is hydrated and reacted with Reactive Green 19 (Sigma-Aldrich, St. Louis, MO), Na_2CO_3 (Sigma-Aldrich, St. Louis, MO), and 20% NaCl (Fisher, Fair Lawn, NJ), incubated for 48 h, and thoroughly rinsed with deionized water. The matrix is equilibrated in ethanolamine (Sigma-Aldrich, St. Louis, MO) for 12 h, rinsed with water, and stored at 4 °C. SEAP is loaded onto the column for 8 h, washed with Tris buffer (Bio-Rad, Hercules, CA), eluted with Na_2HPO_4 buffer (Fisher, Fair Lawn, NJ), and the column is regenerated. The purified SEAP concentration was again measured by the SEAP-activity assay and 200 μg of SEAP per sample was concentrated using 10 kDa centrifugation filters (Waters, Boston, MA) for the permethylation assay. Briefly, SEAP was denatured and digested with trypsin (Promega, Madison, WI) and the glycans were cleaved by N-glycanase (ProZyme, Hayward, CA). The cleaved glycans were purified with Hypersep Hyper Carb SPE cartridges (ThermoFisher, Rockford, IL) using 5% v/v acetonitrile with 0.1% v/v TFA as a wash and 50% acetonitrile with 0.1% v/v TFA to elute the glycans. The elution solution was evaporated under airflow and the glycans were reconstituted and permethylated using methyl iodide in the presence of NaOH and DMSO. The permethylated glycan samples were first cleaned up using liquid-liquid extraction with chloroform and then Sep-Pak PS2 SPE cartridges (Waters, Milford, MA) with elution fractions in 15%, 35%, 50%, and 75% acetonitrile. Eluted fractions were evaporated with a vacuum concentrator, then resuspended in 25 μL of 80% methanol. MALDI-TOF glycan analysis was performed with 10,000 shots at 5000 laser power in positive ion reflector mode with 2,5-dihydroxybenzoic acid matrix using a 4800 MALDI TOF/TOF mass spectrometer (ABSciex, Framingham, MA). The relative glycan percentage was determined as the ratio of the individual glycan peak height to the sum of all glycan peak heights (Supp. Figs. 2, 3).

GnT-II knockdown was performed using transfection of *GnT-II* (CGAAUACCCUGACUCCUUUdTdT) and negative control #1 siRNA (Sigma-Aldrich, St. Louis, MO) using Lonza transfection Cell Line Nucleofector Kit V (Lonza, Basel, Switzerland). *GnT-II* knockdown was confirmed by qRT-PCR (Supp. Fig. 4) using the TaqMan® RNA-to-Ct 1-Step Kit (Applied Biosystems, Foster City, CA), a probe (PrimeTime 5′ 6-FAM/ZEN/3′ IBFQ), and primers (5′-GGGCATTAACGAAGTCCTA GTC-3′; 5′-CAGCTGAATGCTGAATGGAAAG-3′) (IDT, Coralville, IA). qRT-PCR was performed in triplicate on a Cepheid SmartCycler II (Cepheid, Sunnyvale, CA).

3. Theory/calculations

3.1. Glyco-Mapper equations, inputs, and outputs

A draft reconstruction of the glycosylation, metabolism, and nucleotide sugar transport systems was created using the CHO-K1 and Chinese hamster genome annotations, comprehensive reaction databases for candidate enzymatic functions, and published experimental data. To follow the genome reconstruction methodology, glycosylation genes were organized into functional classes, a CHO-specific metabolic map (Supp. Fig. 1) was defined, and substrate and cofactor usage, neutral enzymatic reactions, gene and reaction localization, heteromeric enzyme complexes, isozyme functionalities, intracellular transport mechanisms, and supporting metabolic reactions were all verified. The mathematical model was created using media feed components, fuzzy enzyme parameters $(0-5)$ representing each enzyme's flux to determine the likely glycoform composition (a fuzzy parameter is the quantitative value of a variable that has been transformed to a discretized parameter state within bins) (Sokhansanj et al., 2009), and simplified enzyme reaction kinetics dependent upon the discretized parameter values (that is, a reaction rate that is a single discrete value rather than modeled using, for example, Michaelis-Menton rate expressions). Unbalanced or missing reactions as well as reaction directionality and limitations were identified and rectified. The Glyco-Mapper flux parameters emulate the transcriptomic and proteomic data (if available). If the fluxes are not measured, the parameters are set to minimize the number of differences between the experimental reference glycoform and the resulting Glyco-Mapper glycoform, enabling one to predict the biotherapeutic glycosylation phenotype from single-gene (or multi-gene) changes in a method similar to Khodayari et al. (2014). Specifically, the collection of glycans in production within the cell is initially calculated by evaluating the algorithms demonstrated in Supp. Tables 5–7 per the reactions defined in Supp. Tables 1–2 to model the active glycosylation and CCM gene fluxes. The collection of glycans is then altered using the fuzzy enzyme parameters and simplified enzyme reaction kinetics as listed in Supp. Table 8 to predict a final glycoform.

Glyco-Mapper inputs include the type of recombinant protein (mAb or non-mAb) and a cellular location of the glycoprotein (secreted or intracellula because both parameters affect the potential glycoform. The input list of media sugar components enables nucleotide sugar reaction metabolism calculations and permits the calculation of nucleotide sugar production and availability as additional input variables. Each glycosylation and metabolism gene parameter is an input variable that accounts for the flux of the corresponding enzyme, discretely ranging in value between 0 and 5, to determine the potential glycoform composition using flux profile fitting. Parameter values are adapted from transcriptomic data, but if data is not available, the input parameters were selected empirically in a manual and iterative process in accordance with experimental observations to achieve minimal differences between the predicted and experimental reference glycoforms.

Glyco-Mapper outputs include a list of the predicted glycoform glycan composition as well as a count of the glycans within each glycoform. Glyco-Mapper generates four glycoform lists dependent upon two different parameters, glycan classification (the individual glycans (A3G1, A3G2, A3G3) or glycan monosaccharide groupings (A3G)), and secretion classification (the likely secreted glycoform (glycans to be secreted from the cell) or comprehensive intracellular glycoform composition (glycans that will not be secreted from the cell)). Each combination of parameters yields a slightly different glycoform view and understanding of the glycosylation reaction network. Lastly, an optional user-selected glycan is predicted to be present or absent in the final glycoform, and if absent, the metabolism or glycosylation genes preventing the glycan's production are identified. Glyco-Mapper in the current form is only able to predict each glycan in a binary form (presence or absence), not the various quantitative glycan concentrations or levels. This feature is a result of the lack of traditional enzymatic rate equations (Michaelis-Menten kinetics) and was done with the goal of establishing Glyco-Mapper version 1.0 as a glycosylation network prediction tool that combines ease of use with CHO-specific genomic accuracy.

3.2. Statistical information

For each predicted glycoform, the accuracy, specificity, and sensitivity statistics are solely representative of the predicted glycoforms, and not the reference glycoforms or any combination thereof. The accuracy percentage represents the percentage of correct glycan predictions within the experimentally-modified glycoform and was calculated as the sum of experimentally-validated, present and absent glycans predicted divided by the total number of glycans within the glycoform. The specificity percentage represents the true negative prediction rate, calculated as the number of both predicted and experimentally absent glycans divided by the total number of experimentally absent glycans. The sensitivity percentage represents the true positive prediction rate, calculated as the number of both predicted and experimentally present glycans divided by the total number of experimentally present glycans. The delta accuracy percentage represents the accuracy rate of the glycans that changed either their prediction or experimental status between the reference and predicted glycoforms, calculated as the number of correctly predicted glycans that changed status divided by the total number of glycans that changed status. The relative glycan composition deemed to be statistically significant and a non-minor peak was a composition greater than 1% for all literature and experimental calculations in a method similar to Fukuta et al. (2001).

The *GnT-II* siRNA knockdown qRT-PCR data was analyzed as technical triplicates of biological triplicates (Supp. Fig. 4). The results were statistically analyzed in JMP and the two sample one-sided *t*-test was conducted assuming unequal variances with an alpha of 0.05 and resulting in $p < 0.0001$. The difference between the average *GnT-II* knockdown and negative control samples was 2.51, the t ratio was 7.5, the standard error difference was 0.335, and the degrees of freedom was 15.4.

4. Results

Various cell- or glycoform-engineering strategies (metabolic and glycosylation gene knockouts, knockdowns, or overexpressions) in CHO cell lines have been reported and each of these changes results in an altered recombinant protein glycoform. Between 1999 and 2014, ten publications (Goh et al., 2014; Imai-Nishiya et al., 2007; Kanda et al., 2007; Malphettes et al., 2010; Maszczak-Seneczko et al., 2013; Naso et al., 2010; Onitsuka et al., 2012; Sealover et al., 2013; 2007; Tsukahara et al., 2006; Weikert et al., 1999) describe an engineered change in glycosylation-related gene expression with an accompanying characterization of the resulting glycoform changes from the reference glycoform. These papers collectively altered nine genes affecting eight different nucleotide sugar enzymatic reactions in various combinations among CHO cell lines producing both mAb and non-mAb glycoproteins. The alterations examined in these publications did not extend to the GalNAc or Gal-alpha-Gal portions of the glycoform (448 glycans) that this work is capable of predicting. Thus, the detailed glycoforms are limited to either 156 glycans (non-mAb glycoforms) or 40 glycans (mAb glycoforms). Glyco-Mapper predicted each altered glycoprofile starting with the optimized parameters from the appropriate reference glycoform and then the predicted results were compared against each published glycoform to establish a 96.2% glycan prediction accuracy (1547 of 1608 glycans). The average predicted glycoform sensitivity, specificity, and delta accuracy statistics are 85%, 97%, and 85%, respectively (Table 1). The following sections illustrate the application of the Glyco-Mapper tool towards four different examples of glycoform-engineering strategies from among those reported in the literature and one novel

Table 1
The average accuracy, sensitivity, specificity, and delta accuracy statistics for each cell-engineered glycoform prediction. The average is greater than 80% for all statistics and greater than 95% for both the overall accuracy and specificity. The predictions illustrated in the text are denoted with a (*) and in the Supplemental information with a (ʻ). Genes within parentheses in the following table and figures, e.g. *(Fut8)*, were not genetically altered by the authors of the original publications, but the predicted enzyme flux was altered by one integer to obtain a predicted glycoform more aligned with the experimental glycoform.

Author, Year	Gene(s)	Accuracy	Sensitivity	Specificity	Delta accuracy
Onitsuka et al., 2012*	*ST6Gal1*	92.5%	83.3%	94.1%	50.0%
Goh et al., 2014*	*GnT-I/(Fut8)*	93.6%	78.6%	95.1%	81.8%
Kanda et al., 2007ʻ	*Fut8*	97.5%	100.0%	97.3%	100.0%
Kanda et al., 2007ʻ	*GMDS*	97.5%	100.0%	97.3%	100.0%
Kanda et al., 2007*	*GMDS*/Fuc Feed/*(Fut8)*	97.5%	100.0%	97.3%	100.0%
Maszczak-Seneczko et al., 2013ʻ	*SLC35A3/(GnT-II)*	98.7%	90.0%	99.3%	83.3%
Maszczak-Seneczko et al., 2013ʻ	*β4Galt/(GnT-II)*	98.1%	92.3%	98.6%	85.7%
Maszczak-Seneczko et al., 2013*	*β4Galt/SLC35A3/(GnT-II)*	98.7%	83.3%	100.0%	90.9%
This work	*GnT-II*	94.9%	77.3%	97.8%	75.0%
Malphettes et al., 2010	*Fut8*	92.5%	85.7%	93.9%	87.5%
Tsukahara, 2006	*Fut8*	97.5%	100.0%	97.3%	100.0%
Naso et al., 2010	*SiaA*	97.5%	100.0%	97.3%	100.0%
Sealover et al., 2013	*GnT-I*	92.5%	0.0%	97.4%	100.0%
Imai-Nishiya et al., 2007	*Fut8/GMDS*	97.5%	100.0%	97.3%	100.0%
Weikert et al., 1999	*β4Galt*	95.5%	87.5%	95.9%	0.0%
Weikert et al., 1999	*ST3Gal3*	94.9%	85.7%	95.3%	0.0%
Weikert et al., 1999	*β4Galt/ST3Gal3*	96.2%	81.8%	97.2%	–
Average		**96.2%**	**84.6%**	**97.2%**	**84.7%**

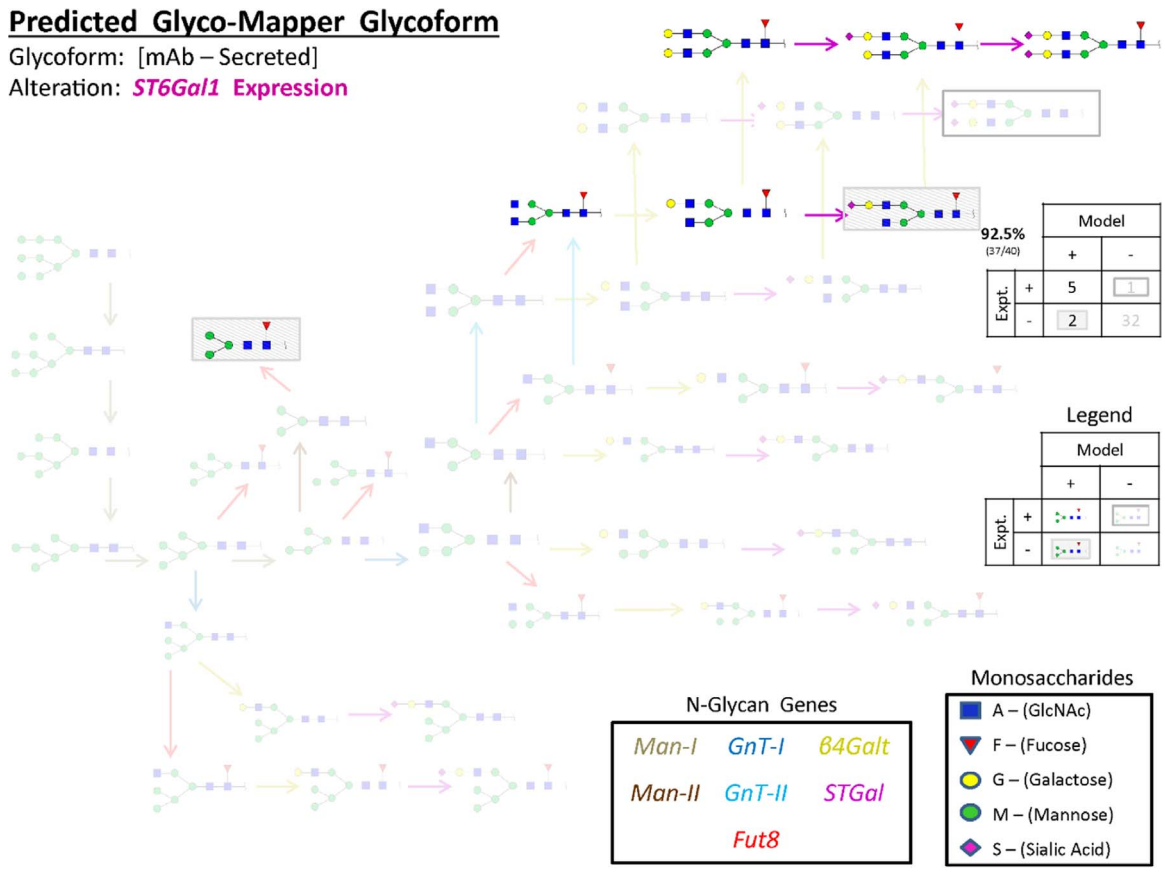

Fig. 1. The Glyco-Mapper prediction of the expression of *ST6Gal1* based on the Onitsuka *et al.* reference glycoform (Supp. Fig. 5). The sialylated glycans FA2G2S1 and FA2G2S2 are correctly predicted to be experimentally present in this gain-of-function glycoform engineering strategy. The majority of glycans (32) were correctly predicted to be absent while five of the six experimentally measured glycans were correctly predicted by Glyco-Mapper to be present, as identified by the "Legend".

glycoform-engineering modification. The fuzzy enzyme parameters from Supp. Tables 1 and 2 and media feed components associated with each of the pictured examples are listed in Supp. Table 9.

The Glyco-Mapper figures (e.g. Supp. Fig. 5, Fig. 1, etc.) are shade-, shape-, and color-coded and the way to decipher each figure is detailed here. The color of each gene within the "N-Glycan Genes" box corresponds with the similarly colored arrows in the figure, detailing the gene responsible for the enzymatic alteration illustrated. The color and shape of each monosaccharide within the "Monosaccharides" box identifies the monosaccharides that compose each of the pictured glycans. The "Legend" identifies each glycan's Glyco-Mapper predicted and reported experimental status. Specifically, glycans predicted by Glyco-Mapper to be present in the final glycoform (Model: + column) are opaque; whereas, glycans predicted by Glyco-Mapper to be absent from

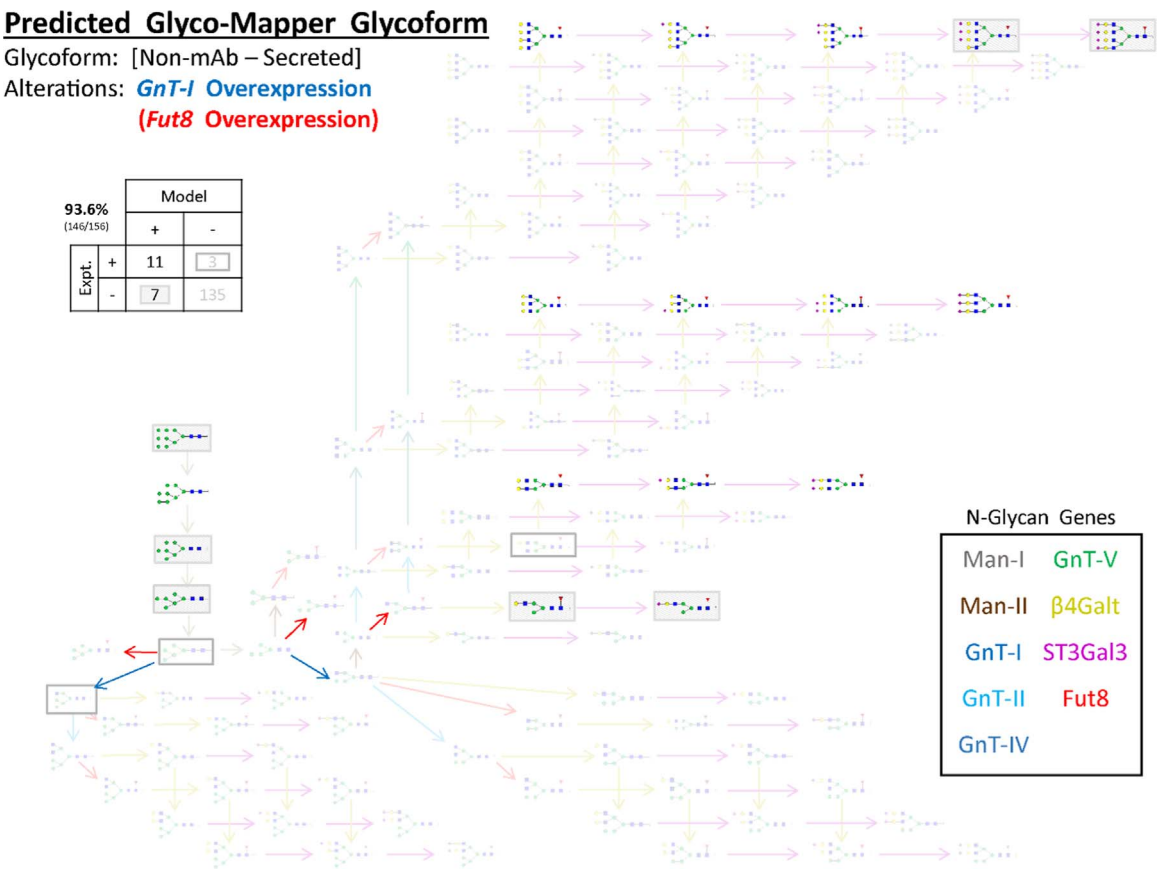

Fig. 2. The Glyco-Mapper prediction of the *GnT-I* overexpression based on the Goh *et al.* reference glycoform (Supp. Fig. 6). The *Fut8* flux was slightly increased. Multiple novel bi-antennary (FA2G2, FA2G2S1, and FA2G2S2), tri-antennary (FA3G3, FA3G3S1, FA3G3S2, and FA3G3S3), and tetra-antennary (FA4G4, FA4G4S1, and FA4G4S2) glycans are all correctly predicted to be experimentally present in this gain-of-function glycoform engineering strategy. The monosaccharide and glycan legends in Fig. 1 are not pictured but still applicable.

the final glycoform (Model: - column) are translucent. Glycans reported to be experimentally present in the final glycoform are pictured as shown in the [Expt. +] row; whereas, glycans reported to be experimentally absent from the final glycoform are pictured as shown in the [Expt. –] row, based on the Glyco-Mapper prediction. Experimental glycan measurements that disagree with the Glyco-Mapper predictions are boxed; whereas, experimental measurements that agree with the Glyco-Mapper predictions are not boxed. The glycan count for each prediction and experimental measurement combination is shown in another box and the accuracy percentage is listed at the top-left corner of the statistics box.

4.1. Strategy 1: Expression of heterologous glycosyltransferases (e.g. ST6Gal1)

Novel glycans may be produced, or a glycoform distribution may be modified, when non-native glycosylation genes are expressed in a host. Glyco-Mapper successfully replicated and predicted the reference and engineered glycan distributions, respectively, as reported by Naso et al. (2010) and Onitsuka et al. (2012). Onitsuka et al. (2012) expressed *ST6Gal1* to increase sialylation, thereby potentially increasing the IgG's biotherapeutic *in vivo* half-life. Glyco-Mapper replicated the wild type glycoform (Supp. Fig. 5) with 39 of 40 correct glycans (3 of 4 present; 36 of 36 absent). When *ST6Gal1* expression was estimated from Onitsuka et al. (2012) (Onitsuka's Fig. 2 and supporting text), the increased *ST6Gal1* flux was modeled in Glyco-Mapper by increasing the *ST6Gal1* parameter from 0 to 3 and Glyco-Mapper predicted (Fig. 1) 37 of 40 glycans correctly (5 of 7 present; 32 of 33 absent). The predicted IgG glycoform resulting from the altered heterologous glycosyltransferase flux was accurate, sensitive, and specific. Supp. Fig. 14 contains screenshots of Glyco-Mapper, detailing how the results shown in Supp. Fig. 5 and Fig. 1 were obtained.

4.2. Strategy 2: Genetic manipulation of glycosyltransferases (e.g. GnT-I)

Genome editing tools are increasingly being used to knockdown, knockout, or overexpress targeted glycosylation genes and alter biotherapeutic glycoforms. Glyco-Mapper successfully replicated studies by Kanda et al. (2007), Weikert et al. (1999), Malphettes et al. (2010), Sealover et al. (2013), Goh et al. (2014), Maszczak-Seneczko et al. (2013), and Tsukahara et al. (2006). In particular, Goh et al. (2014) investigated the effect of *GnT-I* expression in a *GnT-I* knockout CHO cell line with the goal of increasing the sialylation of the glycoprotein erythropoietin (EPO). Glyco-Mapper accurately replicated 149 of 156 glycans (2 of 5 present; 147 of 151 absent) for the wild type (*GnT-I* knockout) glycoform (Supp. Fig. 6); whereas Glyco-Mapper predicted 146 of 156 glycans correctly (11 of 18 present, 135 of 138 absent) when the *GnT-I* overexpression was estimated through an increased *GnT-I* flux (Fig. 2) as well as a slight increase of the *Fut8* flux. The predicted EPO glycoform resulting from the altered glycosyltransferase flux was highly accurate and specific.

4.3. Strategy 3: Genetic manipulation of glycosyltransferase and metabolism genes (e.g. GMDS and Fut8) and Nutrient Feeding Modifications (e.g. Fucose-Feed)

The knockout of a native metabolism gene, or the alteration of a media feed (nutrient composition), can change cellular fluxes and result in a modified glycoform. Two examples of this approach (Kanda et al. (2007) and Imai-Nishiya et al. (2007) were successfully predicted by

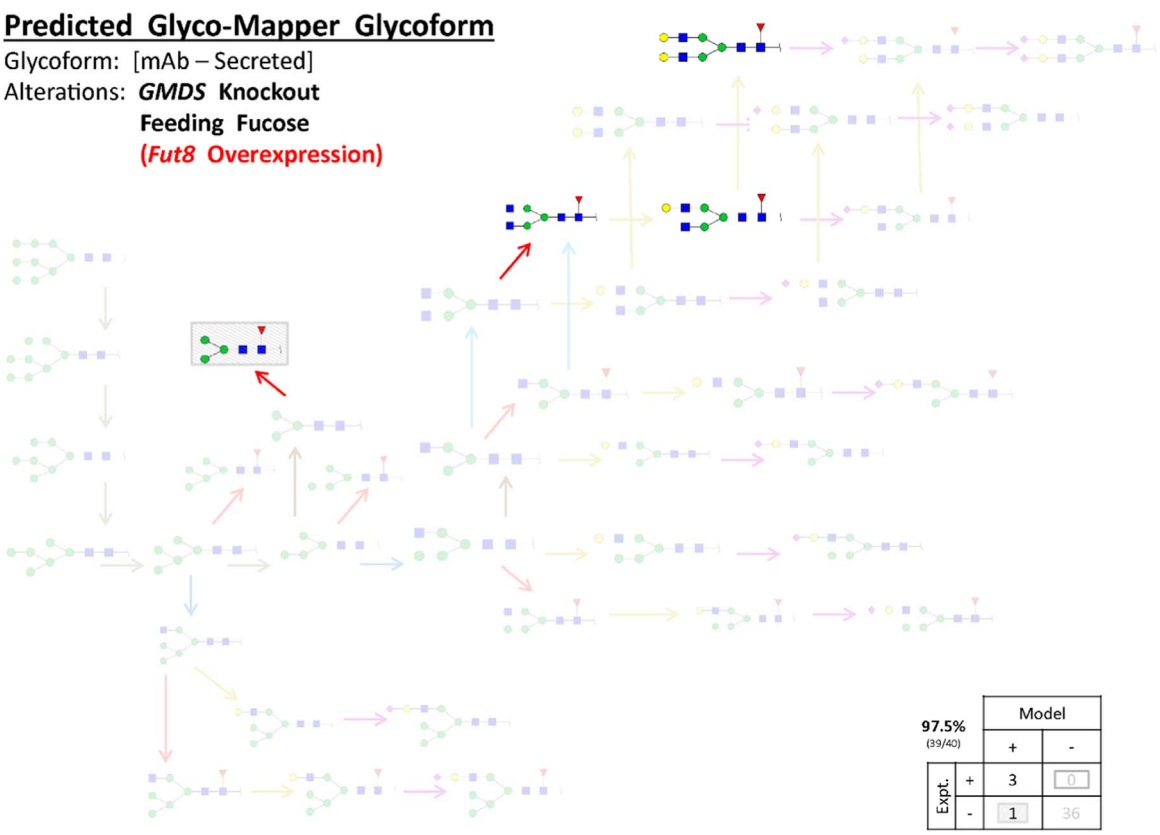

Fig. 3. The Glyco-Mapper prediction of the fucose feeding strategy coupled with the knockout of *GMDS* based on the Kanda *et al.* reference glycoform (Supp. Fig. 7). The *Fut8* flux was slightly increased. The fucosylated bi-antennary glycans FA2, FA2G1, and FA2G2 are correctly predicted to be experimentally present in this metabolic and glycosylation engineering strategy. The N-glycan gene, monosaccharide, and glycan legends in Fig. 1 are not pictured but still applicable.

Glyco-Mapper. Using a mAb (IgG1)-producing CHO cell line, Kanda et al. (2007) independently knocked out *GMDS, Fut8*, and *GMDS* with an altered nutrient-feed containing fucose, thereby affecting the antibody-dependent cellular cytotoxicity through changes in fucosylation (Shinkawa et al., 2003). Glyco-Mapper replicated the wild type glycoform (Supp. Fig. 7) with 37 of 40 correct glycans (5 of 8 present, 32 of 32 absent). When both the *GMDS* knockout (Supp. Fig. 8) and *Fut8* knockout (Supp. Fig. 9) were independently incorporated by forcing the respective fluxes to zero, Glyco-Mapper accurately predicted 39 of 40 glycans correctly (3 of 4 present; 36 of 36 absent). Glyco-Mapper incorporated the nutrient-feed containing fucose and the *GMDS* knockout by forcing the *GMDS* flux to zero (Fig. 3) and accurately predicted 39 of 40 glycans correctly (3 of 4 present; 36 of 36 absent) when accounting for a slight increase of the *Fut8* flux. The predicted IgG1 glycoforms resulting from the modified feeding strategy and the altered metabolic and glycosyltransferase gene fluxes were all highly accurate, sensitive, and specific.

4.4. Strategy 4: Genetic manipulation of glycosyltransferases and nucleotide sugar transporter genes (e.g. SLC35A3 and β4Galt)

Altered nucleotide sugar transport gene fluxes affect the glycoform, whether altered in conjunction with a glycosyltransferase or independently. Glyco-Mapper successfully predicted the glycoprotein glycoforms reported by Maszczak-Seneczko et al. (2013) who knocked down *SLC35A3*, the gene responsible for UDP-GlcNAc transport where reduced transport results in reduced glycoprotein glycan antennarity, and who knocked out *β4Galt*, the genes responsible for the addition of Gal monosaccharides where terminal Gal monosaccharides reduce the biotherapeutic's half-life (Ashwell and Morell, 1974). Glyco-Mapper accurately replicated 153 of 156 glycans (10 of 12 present, 143 of 144 absent) for the wild type glycoform (Supp. Fig. 10); whereas Glyco-Mapper predicted 153 of 156 glycans correctly (12 of 14 present; 141 of 142 absent) when the *β4Galt* knockout was incorporated by forcing the *β4Galt* fluxes to zero (Supp. Fig. 11) and incorporating a slight decrease of the *GnT-II* flux. Glyco-Mapper predicted 154 of 156 glycans correctly (9 of 10 present; 145 of 146 absent) when the *SLC35A3* knockdown was incorporated after decreasing the *SLC35A3* flux (Supp. Fig. 12) from the reference value of 3 (Supp. Figs. 10) to 2, 1, or 0 (the *SLC35A3* flux of 2 prediction is shown in Supp. Fig. 12 but the *SLC35A3* flux of 1 and 0 predictions are not shown as all three predictions are identical) as well as making a minor adjustment to the *GnT-II* flux. Glyco-Mapper accounted for the combined *SLC35A3* knockdown and *β4Galt* knockouts (Fig. 4) accurately by also predicting 154 of 156 glycans correctly (10 of 10 present; 144 of 146 absent) after forcing the *β4Galt* fluxes to zero, decreasing the *SLC35A3* flux in addition to slightly decreasing the *GnT-II* flux. The predicted glycoprotein glycoforms resulting from the altered nucleotide sugar transporter and glycosyltransferase gene fluxes were accurate, sensitive, and specific.

4.5. Novel experimental confirmation using Strategy 2: Genetic manipulation of glycosyltransferases (e.g. GnT-II)

After confirmation of Glyco-Mapper's ability to accurately predict reported changes in literature, Glyco-Mapper's ability to predict a change not previously defined in literature was experimentally tested. The gene *GnT-II* was knocked down using short interfering RNA (siRNA) with the goal of inhibiting bi-antennary glycan formation. Glyco-Mapper accurately replicated 144 of 156 glycans (11 of 14 present, 133 of 142 absent) for the wild type SEAP glycoform (Supp. Fig. 13); whereas Glyco-Mapper predicted 148 of 156 glycans correctly (17 of 20 present; 131 of 136 absent) when the *GnT-II* knockdown was estimated by forcing a *GnT-II* flux decrease (Fig. 5). The predicted SEAP glycoform resulting from the altered flux of the glycosyltransferase

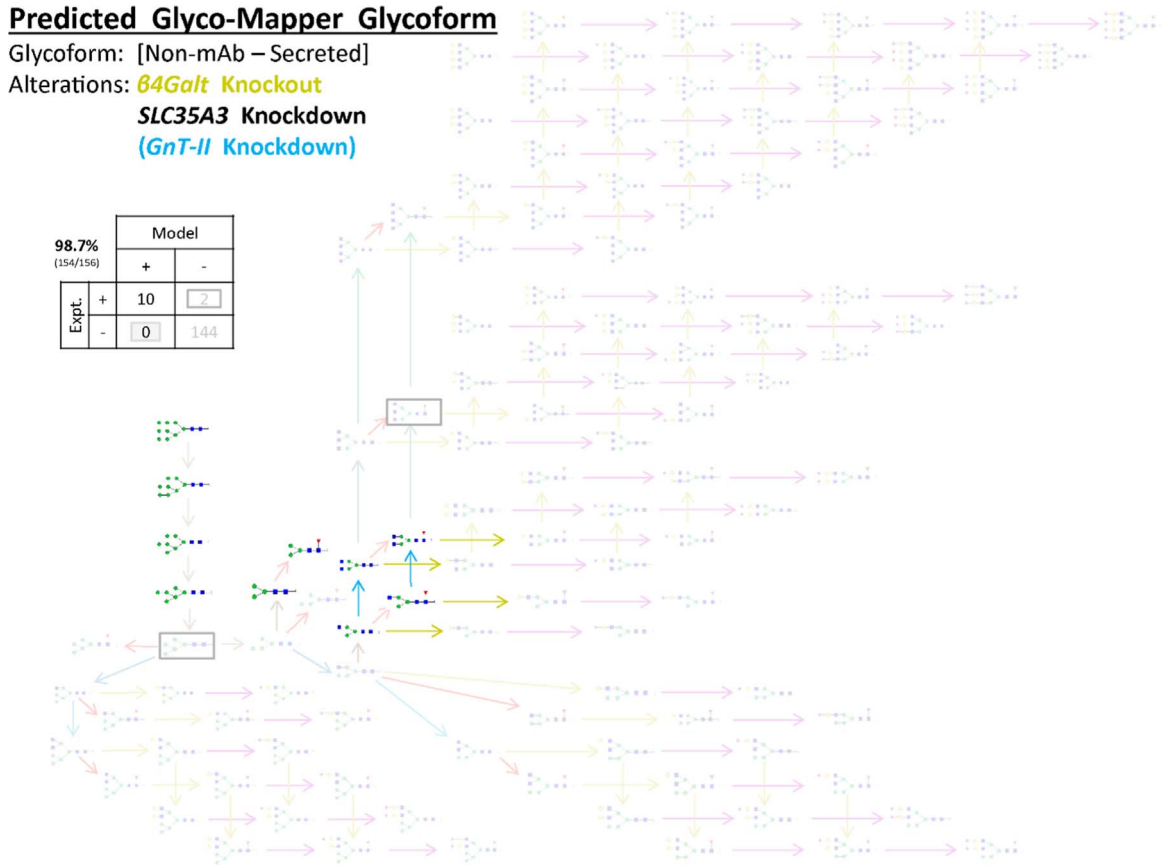

Fig. 4. The Glyco-Mapper prediction of the *SLC35A3* knockdown and *β4Galt* knockout strategy based on the Maszczak et al. reference glycoform (Supp. Fig. 10). The *GnT-II* flux was decreased by one integer. The agalactosylated glycans A2, FA2, A1, and FA1 are all correctly predicted to be experimentally present in this complex glycosylation engineering strategy. The monosaccharide and glycan legends in Fig. 1 and N-glycan gene key in Fig. 2 are not pictured but still applicable.

GnT-II was novel as well as highly accurate and specific.

5. Discussion

The Glyco-Mapper tool developed via the DReaM-zyP method builds upon a protocol for the generation of high-quality genome-scale metabolic reconstructions (Thiele and Palsson, 2010) by incorporating flux flow stoichiometry calculations to predict glycoforms. Manual reconstruction and refinement of the glycosylation reaction network, as well as of the relevant CCM and nucleotide sugar transport and production metabolic pathways, was accomplished using the CHO-K1 and Chinese hamster genome annotations. The incorporation of each media feed sugar and use of discrete, quantized reaction flux parameters for each gene in the reaction network, updated the reconstruction database into Glyco-Mapper, a flux profile modeling tool. Glyco-Mapper provides cell line specific and experimentally relevant discretized glycoform predictions resulting from flux alterations affecting fucosylation, sialylation, galactosylation, antennarity, nucleotide sugar transport, and nucleotide sugar metabolism.

The Glyco-Mapper predicted glycoforms resulting from the overexpression, knockout, and knockdown of glycosylation, nucleotide sugar transporter, and metabolism genes demonstrated an accuracy and specificity of greater than 95% and an average sensitivity and delta accuracy of 85% when compared to experiment. Both engineered mAb glycoforms (nine alterations) and non-mAb glycoforms (eight alterations) were predicted with an average accuracy, sensitivity, and specificity of 96%, 85%, and 97%, respectively (Supp. Table 3). The high predictive delta accuracy and the consistent accuracy, sensitivity, and specificity statistics for both mAb and non-mAb biotherapeutics are indicative of the reliable Glyco-Mapper predictions.

Glyco-Mapper is an experimentally applicable tool that is able to accurately predict cell line-specific glycoforms due to the tool's ability to produce numerous cell line-specific reference glycoforms. These diverse reference glycoforms contain a wide range of unique experimental characteristics, the variety of which are briefly examined in the five reference glycoforms presented in this report. Glyco-Mapper replicated both large and small glycoforms, from 3 glycans (Supp. Fig. 5) to 20 glycans (Supp. Fig. 13). Glyco-Mapper simulated reference glycoforms with glycans localized to both specific and diverse subclasses, from only one bi-antennary subclass (Supp. Fig. 7) to five subclasses (Supp. Fig. 10). Glyco-Mapper replicated reference glycoforms containing glycans that are very similar in structure and also very different in structure, as evidenced by the maximum number of enzymatic reactions between any two measurable glycans in the reference glycoform that ranged from 2 reactions (catalyzed by 1 enzyme) (Supp. Fig. 5) to 15 reactions (catalyzed by 8 different enzymes) (Supp. Fig. 10). The range of these three characteristics depicted by these five reference glycoforms demonstrate Glyco-Mapper's significant versatility. This versatility is currently limited to the binary prediction (presence or absence) of each glycan and future improvements will include the ability to predict the glycan concentration within each glycoform.

Glyco-Mapper is able to successfully predict cell engineered glycoforms using the accurate cell-line specific reference glycoforms. The Glyco-Mapper predicted glycoforms achieve a range of quantized glycoform patterns and all of the glycosylation reaction network fluxes are adjustable, individually and in combination with other genes. Glyco-Mapper predicted quantized glycoform patterns consistent with distinct mAb (Fig. 3) and non-mAb (Supp. Fig. 11) experimental glycoform patterns. Glyco-Mapper predicted significant glycan composition changes resulting from *GnT* and UDP-GlcNAc transporter gene

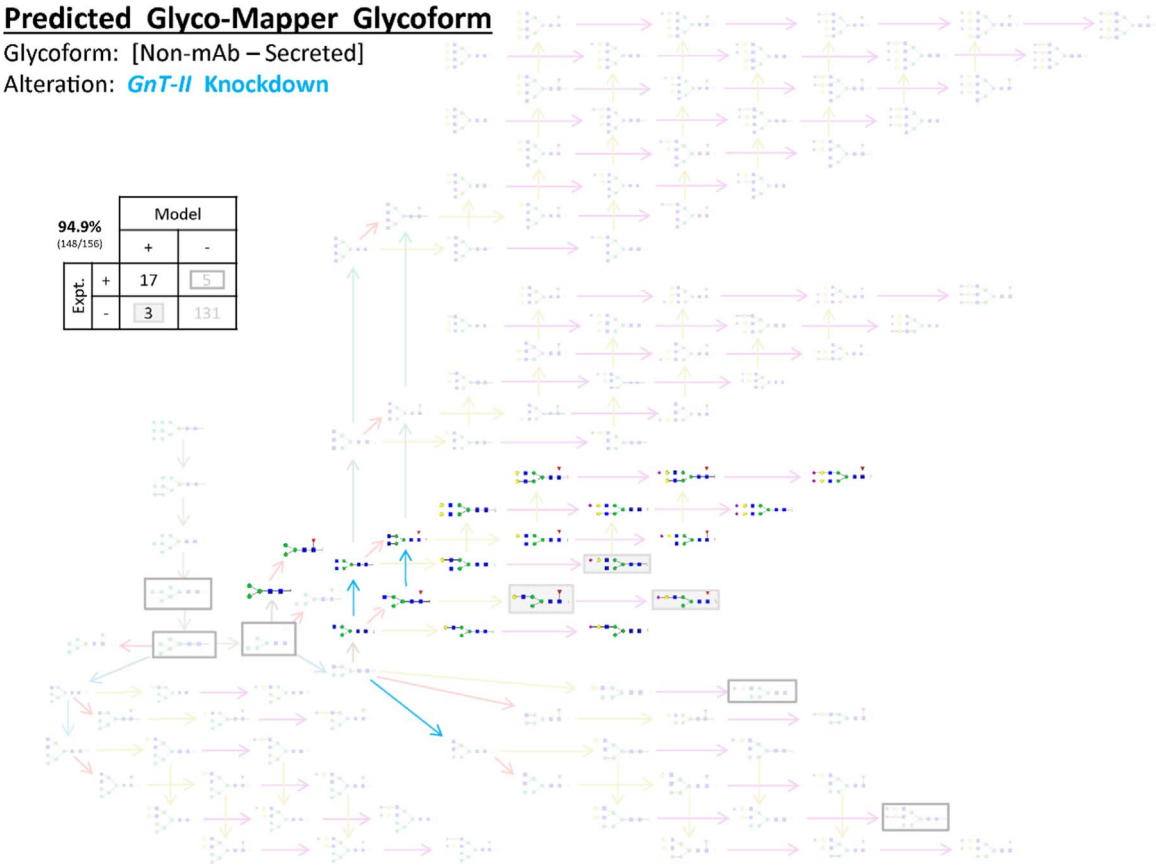

Fig. 5. The Glyco-Mapper predicted *GnT-II* knockdown glycoform is based on the SEAP reference glycoform (Supp. Fig. 13). The single antennae glycans A1, FA1, A1G1, and A1G1S1 are all correctly predicted to be experimentally present in this glycoform engineering strategy. The monosaccharide and glycan legends in Fig. 1 and N-glycan gene key in Fig. 2 are not pictured but still applicable.

alterations (Figs. 1 and 4, respectively) as well as subtle, yet biopharmaceutically-relevant composition changes due to *STGal* and *Fut8* gene alterations (Fig. 1 and Supp. Fig. 11, respectively). The number of predictive glycan changes resulting from an adjusted gene parameter is partially dependent on the reference glycoform as well as on the degree to which gene parameter flux(es) is/are adjusted. While publications typically report engineered glycoform changes, Glyco-Mapper demonstrates an accurate flux representation by predicting a maximum of 15 predicted glycan changes (Fig. 2) as well as a minimum of zero glycan changes (not pictured) for the predictions listed in Table 1. Glyco-Mapper's ability to accurately predict engineered glycoforms enables a wide variety of industrially-relevant glycoforms to be investigated.

Glyco-Mapper predicted glycans may be classified as incorrect for a multitude of reasons, including biological variance within the cellular glycosylation process; variable glycan detection sensitivity of experimental equipment; and unaccounted substrate, product, enzymatic, or other cellular inhibition mechanisms. Specifically, Glyco-Mapper incorrectly predicts glycans in one of two ways: glycans predicted to be measured that were not experimentally observed and glycans predicted to be absent that were experimentally detected. Glyco-Mapper hs incorrectly predicted 24 of 548 glycans (4%) in Figs. 1–5 and the 24 glycans are listed in Supp. Table 4. More than half of these incorrectly predicted glycans are one (9 glycans) or two (4 glycans) active enzyme reactions away from glycans correctly predicted to be present, making most incorrect predictions “slightly off-target”. Other errors may be indicative of unidentified inhibitory factors affecting the final glycoform, as opposed to biological variance or nonsignificant glycan detection. One potential example of an unaccounted factor is the incorrect A2G2S2 prediction in Fig. 1. The reference glycoform was fully fucosylated and after *ST6Gal1* overexpression, the predicted glycoform was also fully fucosylated, yet A2G2S2 was reported and composed a significant percentage (~ 20%) of the experimental glycoform.

Glyco-Mapper can predict the effect of specific gene alterations. It can also help understand the effect of altered fluxes that were not experimentally manipulated or regulated. Within each of Sections 4.2–4.4, one gene flux not manipulated by the authors was slightly altered (by only one integer from the reference glycoform parameter activity) to achieve a more optimal fit between the predicted glycoform and the resulting experimental glycoform. The Fig. 4 prediction (Section 4.4 – Maszczak-Seneczko et al., 2013) decreased the *GnT-II* flux one parameter unit, from 4 to 3. The decreased *GnT-II* activity enabled the production of single-antennae glycans, greatly improving the glycoform prediction accuracy. This observation is consistent with the *GnT-II* flux being impacted by the decreased amount of UDP-GlcNAc available, resulting from the documented knockdown of *SLC35A3*. The Fig. 3 prediction (Section 4.3 – Kanda et al. 2007) increased the *Fut8* flux one parameter unit, from 2 to 3. The increased *Fut8* activity enabled the production of solely fucosylated glycans, greatly improving the glycoform prediction accuracy. This observation is consistent with the *Fut8* flux being altered by the amount of GDP-Fuc available, resulting from the documented knockout of *GMDS* and the feeding of fucose. The Fig. 2 prediction (Section 4.2 – Goh et al., 2004) increased the *Fut8* flux one parameter unit, from 3 to 4. The increased *Fut8* activity enabled the production of solely fucosylated glycans, greatly improving the glycoform prediction accuracy. This discrepancy could indicate that either Glyco-Mapper is not reflective of the correct network (as *Fut8* and *GnT-I* are not linked by multi-enzyme complexes, but both are located within the Golgi) or that there is an unidentified increase in fucosylation activity.

6. Conclusions

Glyco-Mapper predicts glycoforms with high accuracy, specificity, and selectivity after targeted gene manipulations have occurred by combining the power of genome-scale reconstruction with discretized flux flow stoichiometric pathway modeling. Published glycosylation, nucleotide sugar metabolism, and nucleotide sugar transporter gene modifications have all been modeled and the predicted glycoforms averaged a 96% accuracy in specific glycan prediction when compared with experimental results. Upon examination and analysis of the variety of glycoform engineering strategies, the altered glycosylation modifications were correctly predicted regardless of the targeted glycosylation modification or type of biotherapeutic. Glyco-Mapper facilitates an understanding of not only the possible effects, but the likely effects of altering a gene's flux upon the reaction network and the final glycoform through gene knock-outs, knock-downs, or overexpression. The ability to make accurate predictions should enable data-driven selection of beneficial genetic alterations that would be useful to the biopharmaceutical manufacturing community. The DReaM-zyP methodology will enable improved prediction and understanding of biological processes in addition to glycosylation, such as complex signaling networks or the many epigenomic modifications that are currently difficult to mathematically model using published cellular data.

Acknowledgements

We are grateful for support from the National Science Foundation (1412365 and 1247394). We would like to thank Leila Choe for assistance with mass spectrometry data collection, Dr. JongYoun Baik for assistance with figure generation, and Dr. Yu-Sung Wu for discussions and guidance regarding protein purification.

Conflict-of-interest statement

The authors declare no commercial or financial conflict of interest.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2018.03.002.

References

Ahn, W.S., Antoniewicz, M.R., 2012. Towards dynamic metabolic flux analysis in CHO cell cultures. Biotechnol. J. 7, 61–74.
America's Biopharmaceutical Research Companies, 2013. Medicines in development – Biologics – 2013 Report, pp. 1–89.
Ashwell, G., Morell, A.G., 1974. Role of surface carbohydrates in hepatic recognition and transport of circulating glycoproteins. Adv. Enzymol. Relat. Areas Mol. Biol. 41, 99–128.
Bosques, C.J., et al., 2010. Chinese hamster ovary cells can produce galactose-α-1,3-galactose antigens on proteins. Nat. Biotechnol. 28, 1153–1156.
Brinkrolf, K., et al., 2013. Chinese hamster genome sequenced from sorted chromosomes. Nat. Biotechnol. 31, 694–695.
Chen, N., Koumpouras, G.C., Polizzi, K.M., Kontoravdi, C., 2012. Genome-based kinetic modeling of cytosolic glucose metabolism in industrially relevant cell lines: *Saccharomyces cerevisiae* and Chinese hamster ovary cells. Bioprocess Biosyst. Eng. 35, 1023–1033.
Chowdhury, R., Chowdhury, A., Maranas, C.D., 2015. Using gene essentiality and synthetic lethality information to correct yeast and CHO cell genome-scale models. Metabolites 5, 536–570.
Duarte, N.C., et al., 2007. Global reconstruction of the human metabolic network based on genomic and bibliomic data. PNAS 104, 1777–1782.
Fukuta, K., et al., 2001. The widespread effect of β1,4-galactosyltransferase on N-glycan processing. Arch. Biochem. Biophys. 392, 79–86.
Goh, J.S.Y., et al., 2014. Producing recombinant therapeutic glycoproteins with enhanced sialylation using CHO-gmt4 glycosylation mutants. Bioengineered 5, 1–5.
Hammond, S., Kaplarevic, M., Borth, N., Betenbaugh, M., Lee, K.H., 2012. Chinese hamster genome database: an online resource for the CHO community at www.CHOgenome.org. Biotechnol. Bioeng. 109, 1353–1356.
Hayduk, E.J., Lee, K.H., 2005. Cytochalasin D can improve heterologous protein productivity in adherent Chinese hamster ovary cells. Biotechnol. Bioeng. 90, 354–364.
Hills, A.E., Patel, A., Boyd, P., James, D.C., 2001. Metabolic control of recombinant monoclonal antibody N-glycosylation in GS-NS0 cells. Biotechnol. Bioeng. 75, 239–251.
Imai-Nishiya, H., et al., 2007. Double knockdown of a1,6-fucosyltransferase (FUT8) and GDP-mannose 4,6-dehydratase (GMD) in antibody-producing cells: a new strategy for generating fully non-fucosylated therapeutic antibodies with enhanced ADCC. BMC Biotechnol. 7, 84–96.
Jayapal, K.P., Wlaschin, K.F., Hu, W.S., Yap, M.G.S., 2007. Recombinant protein therapeutics from CHO cells – 20 years and counting. Chem. Eng. Prog. 103, 40–47.
Kanda, Y., et al., 2007. Establishment of a GDP-mannose 4,6-dehydratase (GMD) knockout host cell line: a new strategy for generating completely non-fucosylated recombinant therapeutics. J. Biotechnol. 130, 300–310.
Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.
Khodayari, A., Zomorrodi, A.R., Liao, J.C., Maranas, C.D., 2014. A kinetic model of *Escherichia coli* core metabolism satisfying multiple sets of mutant flux data. Metabol. Eng. 24, 50–62.
Krambeck, F.J., Betenbaugh, M.J., 2005. A mathematical model of N-linked glycosylation. Biotechnol. Bioeng. 92 (6), 711–728.
Krambeck, F.J., et al., 2009. A mathematical model to derive N-glycan structures and cellular enzyme activities from mass spectrometric data. Glycobiology 19 (11), 1163–1175.
Kremkow, B.G., Baik, J.Y., MacDonald, M.L., Lee, K.H., 2015. CHOgenome.org 2.0: genome resources and website updates. Biotechnol. J. 10, 931–938.
La Merie Business Intelligence, 2016. Blockbuster biologics 2015. RD Pipeline News (7), 21–28.
Lewis, N.E., et al., 2013. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. Nat. Biotechnol. 31, 759–765.
Liu, G., Neelamegham, S., 2014. A computational framework for the automated construction of glycosylation reaction networks. PLOS One. 9 (6), e100939.
Malphettes, L., et al., 2010. Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. Biotechnol. Bioeng. 130, 300–310.
Maszczak-Seneczko, D., et al., 2013. UDP-N-acetylglucosamine transporter (SLC35A3) regulates biosynthesis of highly branched N-glycans and keratan sulfate. J. Biol. Chem. 288, 21850–21860.
Naso, M.F., Tam, S.H., Scallon, B.J., Raju, T.S., 2010. Engineering host cell lines to reduce terminal sialylation of secreted antibodies. mAbs 2 (5), 519–527.
Onitsuka, M., et al., 2012. Enhancement of sialylation on humanized IgG-like bispecific antibody by overexpression of α-2,6-sialyltransferase derived from Chinese hamster ovary cells. Biotechnol. Prod. Proc. Eng. 94, 69–80.
Ouyang, A., Bennett, P., Zhang, A., Yang, S.T., 2007. Affinity chromatographic separation of secreted alkaline phosphatase and glucoamylase using reactive dyes. Process Biochem. 42, 561–569.
Sealover, N.R., et al., 2013. Engineering Chinese Hamster Ovary (CHO) cells for producing recombinant proteins with simple glycoforms by zinc-finger nuclease (ZFN)-mediated gene knockout of mannosyl (alpha-1,3-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase (Mgat1). J. Biotechnol. 167, 24–32.
Selvarasu, S., et al., 2012. Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture. Biotechnol. Bioeng. 109, 1415–1429.
Selvarasu, S., Karimi, I.A., Ghim, G.H., Lee, D.Y., 2010. Genome-scale modeling and in silico analysis of mouse cell metabolic network. Mol. Biosyst. 6, 152–161.
Shinkawa, T., et al., 2003. The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. J. Biol. Chem. 278, 3466–3473.
Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B.Ø., Ruppin, E., 2008. Network-based prediction of human tissue-specific metabolism. Nat. Biotechnol. 26, 1003–1010.
Spahn, P.N., et al., 2016. A Markov chain model for N-linked protein glycosylation – towards a low parameter tool for model-driven glycoengineering. Metabol. Eng. 33, 52–66.
Sokhansanj, B.A., Datta, S., Hu, X., 2009. Scalable dynamic fuzzy biomolecular network models for large scale biology. Fuzzy Syst. Bioinform. 242, 235–255.
Taniguchi, N., Honke, K., Fukuda, M., 2002. Handbook of Glycosyltransferases and Related Genes. Springer, New York, USA.
Thiele, I., Palsson, B.Ø., 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat. Protoc. 5, 93–121.
Tsukahara, M., et al., 2006. Targeted disruption of a-1,6-fucosyltransferase (FUT8) gene by homologous recombination in Chinese hamster ovary (CHO) cells. Anim. Cell Technol. 175–183.
Umaña, P., Bailey, J.E., 1997. A mathematical model of N-linked glycoform biosynthesis. Biotechnol. Bioeng. 55 (6), 890–908.
US DHHS FDA, 2015. Quality Considerations in Demonstrating Biosimilarity of a Therapeutic Protein Product to a Reference Product, pp. 1–19.
Walsh, G., Jefferis, R., 2006. Post-translational modifications in the context of therapeutic proteins. Nat. Biotechnol. 24, 1241–1252.
Weikert, S., et al., 1999. Engineering Chinese hamster ovary cells to maximize sialic acid content of recombinant glycoproteins. Nat. Biotechnol. 17, 1116–1121.
Xu, X., et al., 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nat. Biotechnol. 29, 735–741.